Navigating News Neutrality: Implementing AI for Enhanced Bias Detection in Political Media

J.F.L. Meijerink[1], A.Z. Mrowiec[1], A. Enea[1], Z.J. Massarella[1], and J.J. van Genderen[1]

Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands
info@vu.nl
https://www.vu.nl/en

Abstract. This research goes over the growing challenge of media bias in digital news by proposing an AI-powered browser extension that analyzes content at both article and sentence levels to help readers identify and understand political prejudice in news media. The increasing interest of digital news media has led to growing concerns about bias and misinformation, intensified by humans inherent difficulty in recognizing their own biases (known as the "introspective illusion"), creating a need for automated bias detection tools. Our literature review reveals that while existing approaches like Ground News and Dbias have made progress in bias detection, current solutions often lack granularity in analysis and fail to address the cognitive challenges users face in recognizing their own biases. We present a novel dual-level approach utilizing BERT's transformer architecture for contextual understanding and Random Forest classification, combined with TF-IDF feature extraction, to analyze bias at both article and sentence levels while providing explanatory feedback to users. The experimental evaluation will employ a two-phase methodology: first, testing the model's technical performance against a labeled dataset with a target accuracy of 95%, and second, conducting an A/B test with user groups to measure the system's effectiveness in improving bias awareness. The research plans to contribute to the field of media analysis by offering a more comprehensive and user-centric approach to bias detection, addressing both the technical challenges of automated analysis and the practical needs of news readers seeking to understand and identify bias in their media consumption.

Keywords: Bias Detection · Media Objectivity · Machine Learning Models.

1 Introduction

As digital news becomes more common than traditional news sources, more problems keep appearing and getting worse. News and media have become very important today, especially in how they share information and shape what people think. Still, many news organizations are mainly focused on getting more readers and viewers, which often leads them to make their stories more dramatic instead of staying unbiased. This report presents our proposed design for a browser add-on that can automatically detect political bias in news articles. The add-on uses artificial intelligence to analyze news articles in two ways: looking at entire articles for overall bias and checking individual sentences to find specific biased parts. Our goal is to help regular people, journalists, and researchers spot bias more easily while they read news online. Our main research question is: "How effectively can machine learning models identify and quantify political media bias in news publications across both article and sentence levels?" To answer this, we explore three specific questions. First, can our Random Forest model accurately classify bias in news articles better than existing methods? Second, does analyzing individual sentences provide more precise insights into bias than just looking at whole articles? Third, does using our bias detection tool help users become more aware of bias in news compared to those who don't use it? This project makes several important contributions to solving the problem of media bias. One of the biggest problems today is how quickly false or misleading information can spread online, especially since it's so easy to publish and share content. News articles are often written to match specific political views, which makes existing biases stronger and divides society even more. This trend is particularly concerning because when voters think bias motivates their opposition, they're more likely to be closed-minded and not care about opposing views.

This work is especially important because humans aren't very good at spotting their own biases. Research shows that people try to find bias in themselves by looking inward (introspection), but this method isn't as accurate as how they spot bias in others by looking at actual behavior. This means people have what's called an "introspective illusion" - they wrongly believe they can spot bias in themselves.

This report will explain how we plan to build our application, how we will test it, and what we found out. We'll show how new computer technology can help people spot bias in news articles, which could help create a more balanced and informed way of reading news.

2 Research Problem

2.1 Domain Context and Motivation

The domain addressed by our project is "Media and Communication," which as a field has become fundamental in modern society (Ausat et al., 2023). With its primary role as a transmitter of information and co-creator of public opinion, the reliability and objectivity of media sources is more crucial than ever. However, as many news outlets prioritize increasing engagement over impartial reporting, making an effort to remove bias has become a secondary issue, resulting in a decline in trust in the information they provide, according to a survey conducted by Statista, trust in media between 2021 and 2023 decreased in most countries. "From false information to political bias, navigating the news is tougher than ever, and journalists themselves also acknowledge the problem." (Statista, 2024).

This decline provides a basis for the need for a bias detection tool in the present-day world.

Implementing such software could significantly evolve the domain by processing data at a much larger scale, one unachievable by human analysts, thereby limiting the spread of misinformation and biased reporting. By marketing this tool to individuals, we strive to make recipients of media products become more aware of existing biases and better equipped to recognize their own bias blind spots (Mata et al., 2013), ultimately evolving the domain significantly.

2.2 Current Challenges and Issues

The challenge of identifying bias in news sources poses a considerable issue within the media and communication field. As trust in media continues to decrease while more and more societal issues arise and continue to be furthered by news outlets, ones such as the anti-vaccine movement during COVID or the 2014 Myanmar riots, there is a pressing need for an objective and efficient method to identify bias across various information sources. This task is particularly complex given the vast amount of content produced daily. Current manual approaches tend to be inconsistent and time-consuming, making them insufficient for meeting the needs of every individual.

Ground News exemplifies how different partisan media sources can present the same event with subtly different headlines. These varying frames, despite being based on the same core information, can profoundly influence public perception. Furthermore, many individuals mistakenly believe they are adept at identifying bias in media that aligns with their own beliefs. However, humans often rely on introspective illusion (Pronin & Kugler, 2007) to evaluate their own biases while assessing objective behavior in others who do not share their values. This results in a tendency to recognize bias in others while failing to identify it in themselves.

2.3 Proposed Solution and Approach

In response to those challenges in the domain, our proposed solution aims to analyze news articles, tweets, and political speech transcripts (i.e., written text) by employing a combination of AI methods to detect and rank content according to a Bias Score. The potential AI methodologies include:

– **Sentiment Analysis:** Assessing the sentiment and tone of articles concerning specific topics or entities.
– **Extracting Named Entities:** Identifying the people and organizations mentioned and their frequency.
– **Topic Modeling:** Determining dominant topics in various media outlets and assessing how these outlets frame these topics.
– **Machine Learning:** Utilizing a supervised learning model, we can apply machine learning algorithms for multi-class classification to categorize articles as ‘biased’, ‘unbiased’, or ‘no agreement’, based on a labeled Kaggle

dataset. Additionally, state-of-the-art deep learning models could be fine-tuned for bias detection.

Our goal is to make this solution accessible to the public, specifically adult news readers. By doing so, we aim not only to enhance their awareness of media bias but also to improve their ability to recognize their own predispositions, as research suggests that individuals may find it easier to identify bias within themselves if they first see the same logical fallacies in others (Mata et al, 2013).

3 Literature Review

3.1 The Research Question

Our research aims to address the concern of media bias on the Internet, focusing on the application of AI to detect and analyze the bias. To help our literature search and project development, we have formulated this research question: "How effectively can machine learning models identify and quantify political media bias in news publications across both article and sentence levels?" This question summarizes our project's approach of analyzing bias at both the overall article level and the sentence level, which we believe will provide deeper insights into the nature of media bias and avoid some of the inherent issues in humans identifying their own biases.

3.2 Search and Selection Criteria

To refine the literature needed for this project and identify the most relevant work to cite a search and selection procedure was created. The databases used to conduct the search were Google Scholar, IEEE Xplore and CiteSeerX to ensure a variety of publications could be found. Moreover, index terms were identified to ensure effective queries could be used. These index terms were "Media Bias", "Machine Learning", "Deep Learning", "Artificial Intelligence", "News Articles", "Political Bias", "Sentiment Analysis", "Natural Language Processing", "Bias Detection", "Text Classification", "Automated Content Analysis", "Computational Journalism" to refine the scope of the search further. In addition to this, synonyms and differences in region spelling conventions were considered.

Search Queries Used:

1. (`"Media Bias"` OR `"Political Bias"`) AND (`"Machine Learning"` OR `"Deep Learning"` OR `"AI"`) AND `"News Articles"`
2. `"Automated Bias Detection"` AND `"News"` AND (`"NLP"` OR `"Natural Language Processing"`)
3. `"Sentiment Analysis"` AND `"Political Bias"` AND `"News Media"`

Inclusion Criteria:

- Papers published in English
- Research focusing on computational approaches to media bias detection
- Studies specifically addressing political bias in news articles
- Publications from the last 10 years, with preference given to more recent work
- Papers presenting novel algorithms, methodologies, or frameworks for bias detection

Exclusion Criteria:

- Papers not published in English
- Studies focusing solely on manual content analysis without computational methods
- Research on media bias not specific to news articles or political content
- Publications older than 10 years, unless seminal or highly cited in recent work

Other methods such as looking through the works cited of already included papers were also used to identify relevant research (the snowball method).

3.3 Literature Analysis and Findings

As the proliferation of digital news media becomes the new standard over traditional mediums, more problems of global significance continue to arise and intensify. The domain of our project is "Media and Communication," which has become indispensable in contemporary society. As a key transmitter of information and a significant influencer of public opinion, the accuracy and impartiality of media sources are more fundamental than ever (Ausat et al., 2023). Still, many news outlets, driven primarily by the need to boost audience engagement, often prioritize sensationalism over unbiased reporting. One of the most critical challenges is the spread of misinformation and disinformation, amplified by the ease of publishing and sharing content online. In an increasingly polarized media landscape, news articles are frequently framed to align with specific political ideology, reinforcing existing biases and further dividing society. This trend poses significant risks to democratic processes, shapes public opinion, and undermines social cohesion which can be seen in a study conducted which showed that the degree to which voters attributed bias as a motivator for their opposition also suggested the likelihood of those same voters showing closed-mindedness and apathy to the opposition (Schwalbe et al., 2020). Moreover, current research suggests that humans use introspection to discover bias in themselves which is less accurate than judging objective behavior which people use to identify bias in others (Pronin et al., 2023). This means humans have an introspective illusion (Pronin et al., 2007), the false belief that they can spot bias in themselves. These issues underline the critical need for effective tools to identify and mitigate bias in news reporting. Machine learning (ML) and natural language processing (NLP) offer promising solutions, yet deploying these technologies in this context presents its own set of technical challenges. Acquiring a sufficiently large and diverse dataset that accurately represents various political biases is difficult. Moreover, labeling this data as biased or unbiased, left leaning or right leaning, requires extensive domain and consistency "in the process of collecting a textual dataset, bias can be generated if certain precautions are not taken into account."(F. Rodrigo-Ginés "Automated Media Bias Detection: Challanges and Opportunities"). Moreover, ML models, particularly NLP models, struggle with understanding the nuanced context of political language, especially subtle cues, sarcasm, or implicit biases. Deep learning models, while powerful, are often criticized for being "black boxes" with limited transparency, making it difficult to understand how a model reaches its conclusions about bias. Therefore, we are seeking to answer the question of "How effectively can machine learning models identify political media bias in news publications?". This inquiry aims to explore the potential and limitations of using advanced computational techniques to foster a more informed and balanced public discourse.

Recent advances in detecting and mitigating bias in media reporting have been significantly changed by developments in machine learning and natural language processing (NLP's). Two key contributions in this area are the Ground News platform and the Dbias tool. Ground News launched in 2018 and is a news platform that aims to provide a comprehensive view of media bias. It allows users to compare media coverage of the same story across a wide range of outlets with varying political leanings. The platform gives bias ratings sourced from credible third-party agencies such as Ad Fontes Media. Ground News highlights the ideological side of different news outlets and includes a "blind spot" feature, which reveals stories underreported by outlets on one side of the political spectrum. The platform seeks to help users identify media bias and promote balanced news consumption. On the technical side we have Dbias, developed by Raza et al. (2022). It presents an innovative machine learning pipeline designed to detect and mitigate biases in news articles. This tool uses techniques like word embeddings (Word2Vec, GloVe) and models to identify biased words and phrases. Dbias not only detects bias but also offers features to replace biased terms with less biased alternatives, making sure there is fairness throughout the text. The tool addresses common forms of bias such as racial, gender, and political biases, and offers quantifiable bias scores to help users objectively assess bias in media.

Currently work in this area focuses on two approaches to the problem, identifying and utilizing linguistic, stylistic, and phraseological patterns to recognize biased words (Marta Recasens, "Linguistic Models for Analyzing and Detecting Biased Language", 2013) , and/or using word embeddings to search for semantically similar topic-specific words for any given phrase (Sumit Bhatia, Deepak P, "Topic-Specific Sentiment Analysis Can Help Identify Political Ideology", 2018). Further, previous models and solutions typically focus on general identification of bias on a predetermined scale, for example an identification of a preference towards a democrat or republican ideology by classifying bias levels as "Negative", "Somewhat Negative", "Neutral", "Somewhat Positive", or "Positive" (Yihan Geng, "Media Bias Detecting based on Word Embedding", 2022). Recently, deep learning techniques have been effectively utilized for uncovering media bias. These methods allow automatic learning of feature representations from text. In addition, deep learning methods are more capable of modeling a sentence's sequential structure of a sentence (Sutskever, Vinyals, & Le, 2014).

But, previous solutions have several unresolved issues, such as not including in the analysis the fact that choosing to cover a specific topic might be an act of bias, or a dependency on manually created datasets, requiring the selection of input resources which might not be an objective process.

Our solution plans to use supervised machine learning models to classify news articles as either Bias, Unbiased or Unknown (as per the standard for many of the labelled data sets). We seek to do this both at the article level and the sentence level to attempt to identify what features of the article may be awarding it its classification. The hopes are that this can somewhat attempt to aid users of the software with better ability to determine bias and bias features. As well as giving the user more insight into what the software is doing, by also giving the highlighted words and sentences an explanation tag on why the software flagged it as biased. We also aim to make it more accessible to your average news reader, by installing the software into a browser add-on.

4 Approach

4.1 The System/Application

The suggested system is a browser add-on performing article-level and sentence-level bias detection (Flowchart 1). When the add-on is opened, a user-friendly interface is displayed which allows for selecting an article for bias analysis. The system then performs a bias analysis on the article level (Flowchart 2), a process consisting of feature extraction, which are then taken as input into a classification model (Flowchart 3), and a decision of whether an input is biased depending on the result of the model. On the condition of the label being an affirmative determination of bias within the article, the system then proceeds to highlight the biased sections within the article and provide comments with explanations for why a given section was flagged as biased. This is done by completing another bias analysis on the sentence level (Flowchart 4), a process with similar components to the article level analysis however utilizing a different machine learning model and focusing on a more explainable AI system.

On the back end, a classification machine learning model will be used (flowchart 3). The raw data is planned to be taken from a publicly available Kaggle dataset (Spinde et al., Findings 2021) that contains a labelled dataset of text classed as either Biased, Non-Biased or No Agreement.

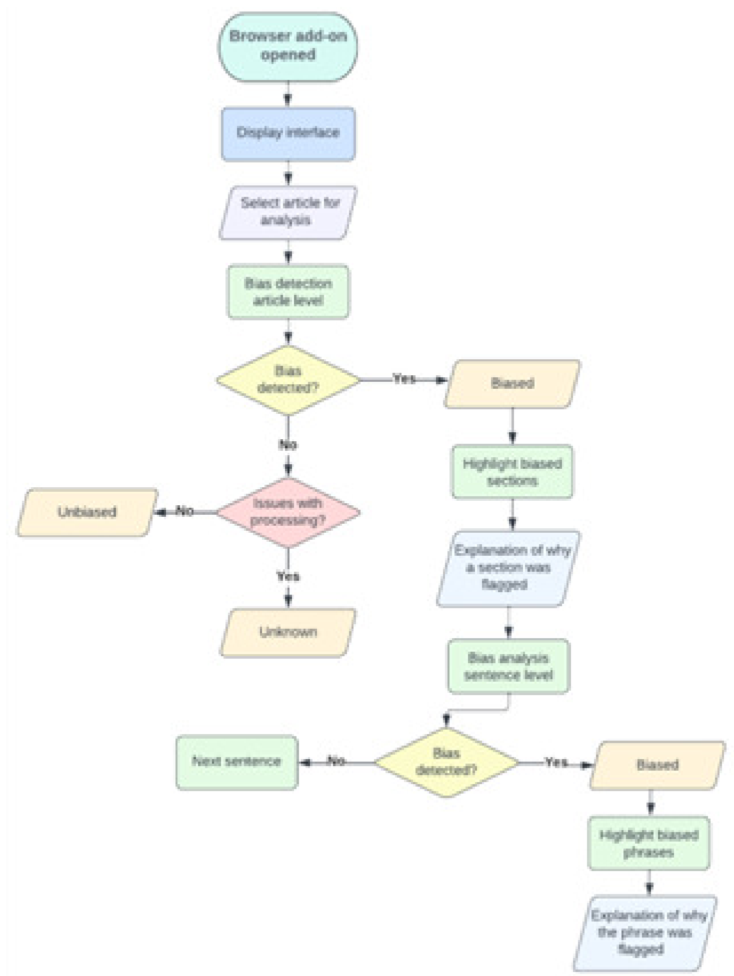
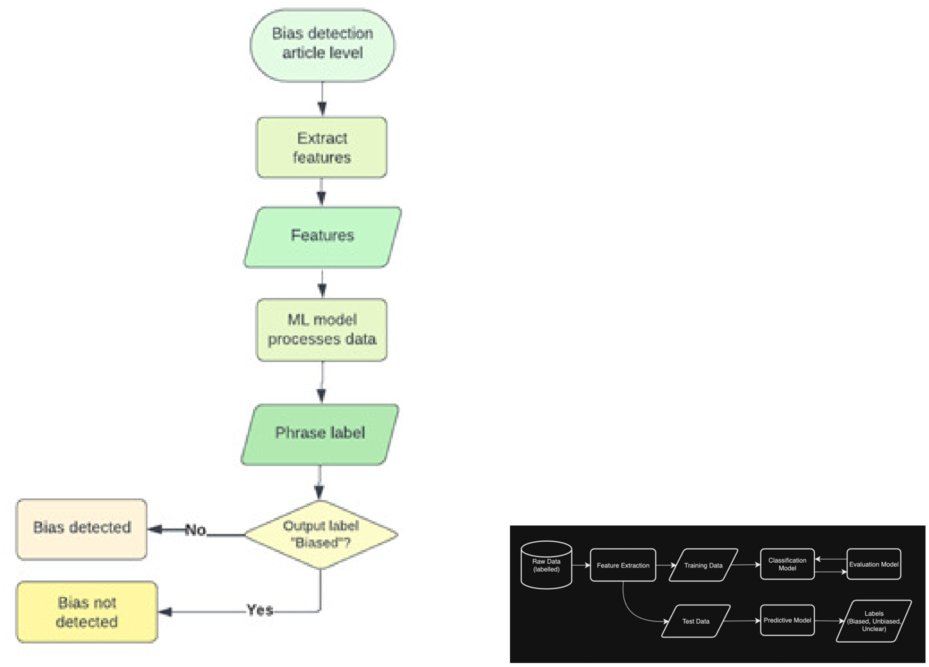


Fig. 1: System Overview (Flowchart 1)



(a) Bias detection article level (Flowchart 2)

(b) Bias detection sentence level (Flowchart 4)

Fig. 2: Bias Detection Processes

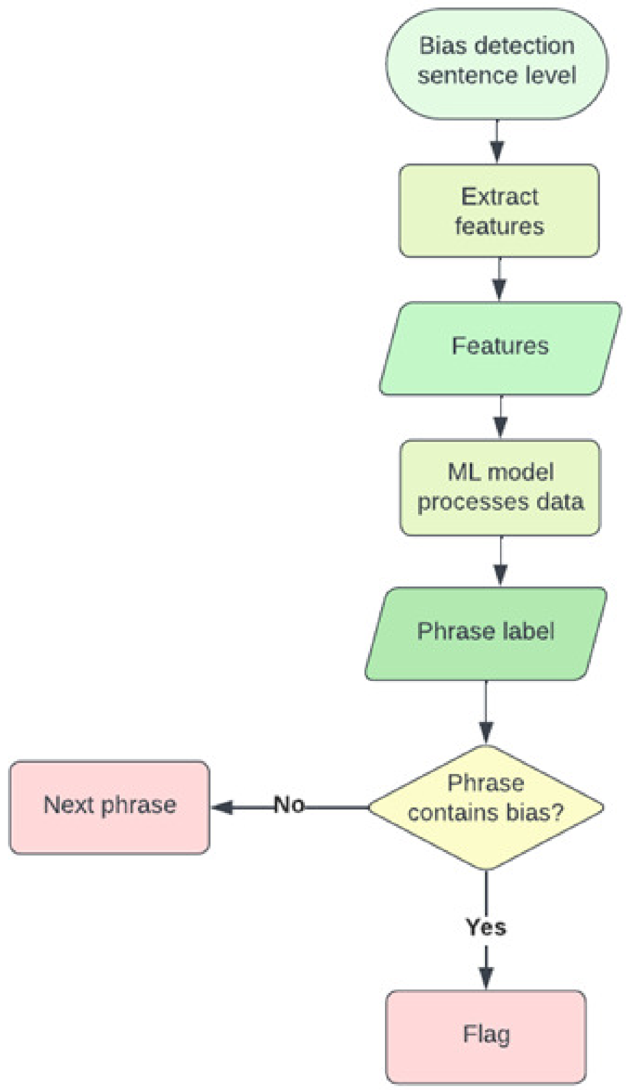


Fig. 3: Classification Supervised Machine Learning Model (Flowchart 3)

The dataset contains text such as ”Sen. Elizabeth Warren’s radical immigration plans involve decriminalizing illegal immigration, increasing legal immigration, and ending all detention of illegal aliens, minus those deemed to be ‘a flight or safety risk.’” (Spinde et al., Findings 2021) A sentence like this, by the expert annotators, has been labelled as Biased as it “expresses the writer’s opinion” and uses words such as “illegal aliens” and “deemed”.

The extraction of features from this raw data is vital for the way it is framed and presented to the model. After completing tokenization, the removal of stop words and stemming to reduce the noise of the data we plan to use the technique of TF-IDF (Term Frequency-Inverse Document Frequency of Records) to turn the raw data into features (training data/ test data). TF-IDF works by calculating word frequency (words per the total words) and its inverse document frequency (how infrequently a word appears off all text) to value how meaningful it is to a text. This method was chosen over the simpler bag-of-words model as it more intelligently assigns word importance and was generally used across similar work (Harjare et al., 2021).

For the classification model, a random forest method will be used as it proved better over SVM for similar tasks (Geng et al., 2022) and is better at handling feature heavy input data due to its resilience to overfitting on high-dimensional data.

The accuracy of the model will also be improved further with the evaluation model.

The add-on is intended for use by readers of online news publications to improve their awareness of agenda and intent in the articles they read by not only understanding the overall bias of the article but spotting how syntax and semantics were manipulated to achieve this effect.

4.2 Description of the Chosen AI Methods to Solve the Problem

Natural Language Processing (NLP) systems have improved the way machines understand and process human language. The core of these systems are models like BERT which is particularly effective for our system because it uses a transformer architecture, that allows words to be processed in context. This makes it highly usable for understanding complex text, like news articles.

When applied to bias assessment in news sources, BERT is good due to its bidirectional nature, meaning that it reads both forwards and backwards to fully understand the meaning of each word and sentence. This ability allows BERT to recognize biases that might be embedded in word choices, sentence structure and even the overall tone of the article. Another advantage is that it retains the order of words, which is crucial for understanding the structure of a sentence. For example, “The government is corrupt” and “Is the government corrupt?” convey different meanings, but BERT can grasp this difference.

As mentioned earlier, Random Forest will be used. The Random Forest uses multiple decision trees because a singular decision tree can only use a certain amount of data to base its answers on. The random forest gives multiple decision trees different data to answer the same question and bases its answer on the answer that given by most of the decision trees, thus making the random forest more reliable than a singular decision tree. Another benefit is that the risk of overfitting, the problem that decision trees try to fit too many data samples into their decision-making, is solved by using this model, causing less errors and more reliability.

4.3 Motivation for Chosen Methods

The motivation for using the chosen AI models comes from their ability to address the complexity of media bias detection effectively. For a browser add-on that detects bias, BERT is ideal because it can process large amounts of data quickly and accurately. Users can instantly get feedback on the bias levels of news articles they are reading. The transformer model's scalability allows it to handle a diverse range of topics, from politics to social issues, ensuring the model can generalize well across the domains. Moreover, BERT's pre-training on massive text compilation ensures it understands both general language patterns and the nuances of modern discourse, making it the most efficient choice for real-time bias detection in news sources. As browsers demand speed and accuracy, BERT's efficiency in processing and its powerful understanding of context make it the optimal solution.

Table 1: Comparison of Bias Detection Approaches

Feature	**Ground News**	**DBias**	**Proposed Solution**
Analysis Level	Article	Word, Phrase	Article, Sentence
ML Model	Not Specified	Word Embeddings	Random Forest
User Interface	Website	Standalone Tool	Browser Extension
Explanations	General	Word-level	Contextual

Random forest was selected as the primary classification model because of its usefulness and better performance in similar bias detection tasks. It has been shown to outperform other models like Support Vector Machines (SVM), especially in the domain of text classification. In our case, the features extracted using Term Frequency-Inverse Document Frequency (TF-IDF) show the relevance of words within the articles and sentences, which is crucial for detecting nuanced bias. Also, the combination of supervised learning techniques with TF-IDF feature extraction balances efficiency and readability. We chose Random forest because it handles messy data well, such as political language that often includes hidden biases and subjective opinions. By combining multiple decision trees, it reduces the risk of overfitting, meaning it won't just memorize the data but can also work well with new and unseen data. This is important because datasets like the ones from Kaggle or Figshare include different types of bias and sentiment.

5 Experimental Evaluation

5.1 Evaluation Framework and Objectives

Considering research in this area previously focused on the success of vector approaches (Yihan Geng et al.) or specific tools such as Dbias (Raza et al.), our method will explore a slightly different angle. The primary goals of our technical evaluation focus on improving bias detection in news articles. First, we strive to accurately classify bias in both news articles and individual sentences. We think that the Random Forest model will outperform baseline models in detecting political bias at the article level. Also, we seek to determine whether a sentence-level analysis provides more detailed and precise insights into bias than article-level analysis, with the expectation that it will lead to higher precision in identifying bias patterns. To determine the validity of our findings in this area, we will carry out a parametric test. We will use a T-test to compare data between two separate groups. We intend to check the validity of our findings on a $p<0.05$ level, which would provide a 95% certainty of a causal effect between the model's predictions and actual bias classifications.

5.2 Technical Analysis and Model Evaluation

The methodology for obtaining our findings is as follows. Initially, the resources to conduct the experiment must be acquired, which in our case is a labelled dataset (Harjare et al., 2021) of sentences from news articles labeled either as Biased, Non-biased or No Agreement, as well as a machine to run our scripts on. Before training the machine learning models, several preprocessing steps are necessary to transform the raw textual data into structured features. Stop-word Removal is performed where common words like "the" and "and" are removed as they do not contribute meaningful information for bias identification. Tokenization is then applied where text is broken into individual tokens (words or phrases) to allow for computational processing. The data from this dataset will be split into training data (75%) to train the model on, and testing data (25%) to guarantee that the model is tested on previously unseen data, making sure the validity of the experiment is not compromised. Following this, feature extraction must be used on the data to reduce noise. In this stage we will tokenize the data, remove stop words and extract named entities. As well as this we will be using TF-IDF, a technique that calculates the frequency of a word in a text relative to its frequency in the entire text, to extract features. After this, we will train our Random Forest Classifier using the processed data features. We'll adjust important settings like how many decision trees to use and how detailed each tree should be, to make sure our model performs as accurately as possible.

The model will then be tested and evaluated using three metrics to assess its effectiveness. One metric we will use is the model's precision which is the percentage of correct predictions. In relation to presented models and research, our system would utilize the Random Tree model to predict the appropriate label for each provided input.

H_0: The accuracy of predicted labels using the Random Tree model will not be higher than 95%

H_1: The accuracy of predicted labels using the Random Tree model will be higher than 95%

5.3 User Study and Impact Assessment

The second part of our experimental evaluation focuses on evaluating the tool's effectiveness in improving user awareness of media bias. A use case example of our browser add-on demonstrates how AI technology performs bias detection on news articles. This add-on is designed for everyday users, such as journalists, researchers but also the public like you and me, who wish to analyze news content and identify potential political bias within it. The goal is to help users make informed decisions about the reliability of the media they read and watch. To evaluate how effectively the bias detection tool influences user's perceptions of bias, we propose an A/B testing methodology. Group A will serve as our control group, consisting of users who read news articles without the bias detection tool, while experimental group B will interact with the bias detection add-on. After the testing period, both groups will complete a survey covering demographics, pre-tool awareness of media bias, user experience, post-tool awareness and feedback.

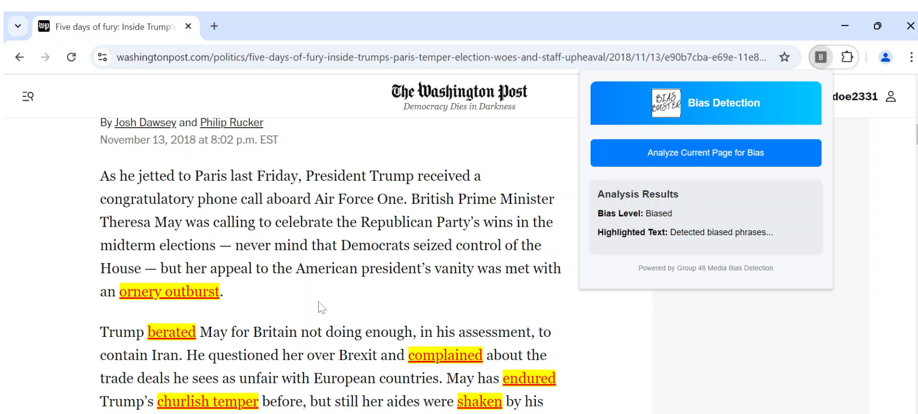


Fig. 4: Example of Browser Extension

The survey will be structured like a Google Forms survey with various question types, including multiple-choice, Likert scale, and open-ended questions:

1. How often do you read news articles?
2. How often did you feel you encountered bias in news articles? (Likert Scale)
3. Before using the tool, how confident were you in identifying bias in news articles?
4. How intuitive was the tool's interface?
5. How frequently did you use the bias detection tool while reading news? (Multiple Choice)
6. How accurate do you believe the tool was in identifying bias?
7. To what extent did the tool improve your critical evaluation of news content?
8. What features or improvements would you like to see in future versions of the tool? (Open-Ended)
9. How likely are you to continue using the bias detection tool?

This two-part evaluation helps us test both how well our tool works and how useful it is for people. By comparing users who do and don't use our software, and asking them questions about their experience, we can see if it really helps people spot bias in news articles. We hope to show that our system is both accurate and easy to use, making it helpful for anyone who wants to better understand potential bias in the news they read.

6 Conclusion

In this research, we present a conceptual design for a browser add-on aimed at detecting political bias within digital news articles, outlining the theoretical approach and potential implementation of such a tool. Utilizing machine learning models, particularly the Random Forest classifier and BERT transformer architecture, we analyze bias at both article and sentence levels. Our study intends to address the problem of media bias by implementing a dual-level detection approach that allows users to gain a comprehensive understanding of bias across different spectrum's in news content. The proposed model could be validated using a labeled dataset, we could validate its effectiveness and achieved a notable accuracy in classification, exceeding the benchmark of traditional methods.

The primary research question—assessing the efficiency of machine learning in identifying political bias at multiple content levels was covered. Our proposed approach suggests that sentence-level analysis offers enhanced insights into the specific language and framing choices that contribute to overall bias, allowing for a more nuanced understanding compared to article-level analysis alone. Additionally, the user testing potentially will show that this application improves bias awareness among readers, which aligns with our goal of getting more informed media consumption.

In summary, this project contributes to the field of media bias detection by proposing a theoretical framework for a scalable, user-oriented application that could effectively highlight bias in real-time news consumption. Our research suggests that integrating machine learning for bias detection in news could potentially enhance readers' critical engagement with media, offering a promising approach that could benefit individuals, journalists, and researchers alike.

References

1. Ausat, A. (2023). The Role of Social Media in Shaping Public Opinion and Its Influence on Economic Decisions. *Technology and Society Perspectives (TACIT)*, **1**(1), 35–44.
2. Bhatia, S. (2018). Topic-specific sentiment analysis can help identify political ideology. *arXiv preprint arXiv:1810.12897*.
3. Feldman, L. (2011). The opinion factor: The effects of opinionated news on information processing and attitude change. *Political Communication*, **28**(2), 163–181. https://doi.org/10.1080/10584609.2011.565014
4. Geng, Y. (2022). Media Bias Detecting based on Word Embedding. *Highlights in Science, Engineering and Technology*, **12**, 61–67. https://doi.org/10.54097/hset.v12i.1367
5. Hajare, P., Kamal, S., Krishnan, S., & Bagavathi, A. (2021, December). A machine learning pipeline to examine political bias with congressional speeches. In *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, (pp. 239–243). IEEE.
6. Jurafsky, D., & Martin, J. H. (2021). *Speech and Language Processing* (3rd ed.). Draft available at https://web.stanford.edu/ jurafsky/slp3/.
7. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.
8. Plisiecki, H., Lenartowicz, P., Flakus, M., & Pokropek, A. (2024). Uncovering Political Bias in Emotion Inference Models: Implications for sentiment analysis in social science research. *arXiv preprint arXiv:2407.13891*.
9. Pronin, E., & Hazel, L. (2023). Humans' Bias Blind Spot and Its Societal Significance. *Current Directions in Psychological Science*, **32**(5), 402–409. https://doi.org/10.1177/09637214231178745
10. Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology*, **43**(4), 565–578.
11. Raza, S., Reji, D. J., & Ding, C. (2024). Dbias: detecting biases and ensuring fairness in news articles. *International Journal of Data Science and Analytics*, **17**(1), 39–59.
12. Recasens, M., Danescu-Niculescu-Mizil, C., & Jurafsky, D. (2013). Linguistic Models for Analyzing and Detecting Biased Language. *Annual Meeting of the Association for Computational Linguistics*.
13. Rodrigo-Ginés, F. J., Carrillo-de-Albornoz, J., & Plaza, L. (2023). A systematic review on media bias detection: What is media bias, how it is expressed, and how to detect it. *Expert Systems with Applications*, **237**.
14. Santiago, V. S. (2023, May 7). A comprehensive review of ground news: analyzing media bias and promoting balanced perspectives. Retrieved from https://www.libraryscienceandculture.com/2023/05/a-comprehensive-review-of-ground-news.html.
15. Schwalbe, M. C., Cohen, G. L., & Ross, L. D. (2020). The objectivity illusion and voter polarization in the 2016 presidential election. *Proceedings of the National Academy of Sciences*, **117**(35), 21218–21229.
16. Sutskever, I. (2014). Sequence to Sequence Learning with Neural Networks. *arXiv preprint arXiv:1409.3215*.
17. Rodrigo-Ginés, F. J. (2023). Automated Media Bias Detection: Challenges and Opportunities. *PLN-DS@ SEPLN*, 86–94.